Applications of statistical methods and techniques to auditing and accounting

van Batenburg, P.C.; Kriens, J.

SITY
RSITEIT
BRABANT
POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

[redacted]

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

APPLICATIONS OF STATISTICAL METHODS AND TECHNIQUES TO AUDITING AND ACCOUNTING

Paul C. van Batenburg and J. Kriens

FEW 493

Applications of statistical methods and techniques to auditing and accounting

Paul C.van Batenburg and J.Kriens

Touche Ross Nederland
Center for Quantitative methods and Statistics

Anthony Fokkerweg 1
NL 1059 CM Amsterdam
The Netherlands

Invited paper for the 48th Conference of
the International Statistical Institute,
Cairo, Egypt, 9-17 September 1991

Applications of statistical methods and techniques to auditing and accounting

Paul C.van Batenburg and J.Kriens

Abstract

In this paper we review a number of statistical methods and techniques that are applied in auditing. First, a brief introduction is made into what auditing means, how it works, and when statistics comes in. In the subsequent sections we will go in to:

- testing and estimation procedures, leading to decision rules used by auditors to determine their sample sizes, and leading to techniques to estimate the population quality from the sample results;
- methods to ascertain a stipulated quality of a population if errors found during the statistical procedure can be corrected;
- the methodological question how to use information about the population quality that is gained preliminary to the drawing of the sample. In this section, we present a Bayesian model to overcome the well-known drawbacks of the audit assurance model which is used by many auditors.

Sommaire

Dans ce papier nous nous proposons de passer en revue un certain nombre de méthodes et de techniques statistiques que l'on applique dans la science de l'audit. Dans une introduction sommaire nous nous arrêterons d'abord sur ce que signifie l'audit, sur son functionnement, et sur les possibilités d'application de la statistique. Dans les sections suivant nous traiterons:

- des procédures de test et d'estimation qui mènent à des regles de décision utilisées par les auditeurs pour déterminer l'effectif de leurs échantillons et qui mènent également à des techniques pour estimer la qualité de la population-parent à partir des résultats de l'échantilon;
- des méthodes pour assurer une qualité déterminée d'une population-parent lorsque des erreurs découvertes pendant la procédure statistique peuvent être corrigées;
- la question méthodologique du comment utiliser l'information sur la qualité de la population-parent obtenue avant le tirage de l'échantillon. Nous présentons un modèle Bayesien pour éviter les désavantages bien connus de l'"audit assurance model" utilisé par tant d'auditeurs.

1. Introduction

Most corporations and many non-profit organizations yearly produce financial statements consisting of a statement of profits and losses and a balance sheet of assets and liabilities. The first contains entries like e.g. sales, costs of sales, selling expenses, income taxes, the second entries like cash, receivables, inventories and buildings as assets, and loans payable, accounts payable and stockholders as liabilities. These accounts are made for the sake of parties interested in earnings, financial position and continuity of corporations and non-profit organizations; one may think of shareholders, government agencies, tax officials, and so on.

In many countries these financial reports are considered to be of so much importance that it is legally required not only to publish them, but also to have these financial statements examined on their reliability by an impartial and independent expert. This work is being carried out by auditors.

It is the auditor's task to give a statement that a firm's balance sheet and profits- and losses account give 'a true and fair view of profits and losses within the clients book year, and of the client's assets and liabilities at the end of that year'. (Translated from the officially prescribed phrases for Dutch auditors; in other countries different, allied, formulations are used). To enable the auditor to make such a statement, he should obtain sufficient appropriate audit evidence.

The way an auditor obtains this evidence is rather complicated. He does not just check directly the data mentioned in the yearly financial reports, if that were possible at all. Once an auditor accepts the engagement to perform an audit, an audit process is carried out according to an audit plan laid out in several phases. To get a rough idea we may paraphrase this process as follows.

1. In the first phase the audit plan is developed. This preliminary phase deals with the decision how to perform the audit. The preparation of the audit strategy is based on a study of structural and procedural measures of the firm's <u>a</u>ccounting <u>o</u>rganization (AO for short) and <u>i</u>nternal <u>c</u>ontrol (IC) procedures which are part of the accounting system.

    The AO has the primary responsibility for supplying reliable information in all aspects; it has the potentialities to correct errors during accounting handling. Management is responsible for implementation and execution of IC.

2. This is followed by an evaluation and testing of the proper working of the internal controls. This phase is generally called 'compliance testing'.

3. The third phase, called 'tests of transactions', is aimed at the data resulting from operational activities (stream of transactions) and other data-files from which the financial statements are prepared. In this phase it should be determined that the financial statements reflect the necessary information correctly and clearly.

The auditor's evaluation of a firm's AO and IC is therefore a very important part of the audit process. This evaluation will often induce the auditor to give recommendations onto how to improve aspects of AO/IC. This can eventually lead to the idea that the main task of the auditor is not just to evaluate the reliability of the financial statements, but to assist an organization in such a way that by continually improving AO/IC a situation is created in which reliable financial statements will be produced almost automatically. In this respect, the tasks of auditors may either be more concentrated on yearly reports or more on AO/IC (depending on national traditions, type of clients, and so on), but 'an audit' always implies more than just evaluating yearly financial statements.

In view of such a broad commitment it is not amazing that there are rich opportunities to apply statistical sampling methods. Since about the beginning of the 1960's, this is really being done; not only in theoretical articles but also in practice. However, many statisticians may be very surprised about the statistical methods used. Besides sound procedures, many questionable practices are applied.

As it is not possible to discuss all existing problems and methods used we have to restrict ourselves to a few of them. We will shortly review:

- (some) testing and estimation procedures (section 2);
- the Average Outgoing Quality Limit method (section 3);
- the question how to integrate all sources of available audit information (section 4).

2. Testing and estimation procedures

2.1. Nomenclature

A random sample of n items is taken from a population consisting of T monetary units divided over N strata (invoices, wage slips, entries in a computer file, etc.):

- in a monetary unit sample **(MUS)** these n items are taken from the population defined as the sequence 1,2,...,T. Each monetary unit is of equal importance to the audit, and (/because it) is supposed to have the same probability of being an error.
  We emphasize that MUS samples are not equivalent to PPS (probability proportional to size)- samples, as is often stated, for example by Cox and Snell (1979) and Godambe and Thompson (1988). Contrary to PPS-sampling, in MUS-sampling an invoice (stratum) can be sampled more than once because more than one monetary units

belonging to that invoice have been drawn. Moreover, in a population monetary units may be present that do not belong to an existing invoice;
- in a 'strata unit sample' **(SUS)** these n items are taken from the population defined as the sequence 1,2,...,N. Each stratum is of equal importance to the audit, and (/because it) is supposed to have the same probability of being an error.

The fraction of errors in a population, either as a fraction of T, or as a fraction of N, is denoted by p. The number of errors in the sample is $\underline{k}$ (random variables will be underlined in this paper), which follows a hypergeometric distribution. Often, this probability distribution is approximated by a binomial or a Poisson distribution.

2.2 Hypothesis testing

The most general formulation of the statistical test is:

$$H_o : p \le p_o \text{ against } H_1 : p > p_1.$$

H_o is rejected if k exceeds a critical value k_1 and H_1 is rejected if k does not exceed a value k_o. If k falls in between, the sample size is increased, and new values k_o and k_1 are calculated. This test can be seen as the most informative method: the auditor is able to test his ideas on the required quality of the population, under the restrictions that both the type I and type II errors are below a chosen critical level.

Practical drawbacks of more-stage and sequential sampling procedures in an auditing context are mentioned in Kriens (1979). As a consequence of these drawbacks, most auditors prefer fixed-sample size methods. Moreover, they prefer to choose identical values for p_o and p_1, and test the set:

$$H_o : p \ge p_o \text{ against } H_1 : p < p_o.$$

This particular set of hypotheses is taken, instead of its reverse, because now the type I error implies rejecting $p \ge p_o$ when in fact $p \ge p_o$ holds, so wrongly accepting a population. This is, of course, much more important to the auditor than not rejecting $p \ge p_o$ when $p < p_o$ holds: wrongly rejecting the population.
Sample sizes are calculated for different values of k_o and α_o from:

$$P[\underline{k} \le k_o \mid N \text{ or } T,n,p_o] \le \alpha_o.$$

How auditors choose α_o will be the subject of section 4.

In many sampling manuals used by audit firms, a chart like chart 2.1 is available, to be used in monetary unit samples in the following manner:

1. determine the critical error for the population to be audited, p_o as a percentage or $p_o T$ as an amount;
2. determine the presumed error amount $p'T$, the amount of errors that is expected to be present in the population;
3. compute $p'T$ as a percentage of $p_o T$;
4. read k from the chart in the row corresponding with a percentage that is at least equal to the result of step 3;
5. follow that row to the sample size correponding to the critical error rate p_o.

If the sample of the size determined as described above yields a number of errors equal to k, the following conclusions can be drawn:

- the Poisson upper 95%-confidence limit for the amount of errors in the population equals the critical amount $p_o T$;
- the best estimate for the amount of errors in the population equals $(k/n)T$, so, at least the presumed error amount $p'T$.

Chart 2.1: sample sizes and acceptance numbers for different values of critical error rate and presumed error (α_o=0.05).

critical value k	upper limit (α_o=0.05)	presumed error as % of $p_o T$	critical error rate p_o=1%	2%	5%	10%
			sample size n			
0	3.00	0%	300	150	60	30
1	4.75	≤ 21%	475	238	95	48
2	6.30	≤ 32%	630	315	126	63
3	7.76	≤ 39%	776	388	156	78
4	9.16	≤ 44%	916	458	184	92
5	10.52	≤ 48%	1052	526	211	106

An example: population size 1 million, α_o=0.05 and p_o=2%, presumed amount of errors 8000. This amount equals 40% of the critical amount of errors. The row '≤ 44%' and the colomn '2%' of chart 2.1 intersect at k=4 and n= 458.
If, in a sample of size 458, 4 errors will occur, the 95%-upper limit for the population error fraction will be 916/458= 2%, and the 'best estimate' will be 4/458 of 1 million = 8734.

As statisticians, we think this method is not as straightforward as it is presented. First, the problem is a problem of hypothesis testing, but it is not explicitely formulated as such. Second, we have doubts about the use of a 'presumed error', determining the acceptance number k. Third and perhaps more important, is the consequent neglect of the fact that though the type II error has implicitely been weighed against sampling costs, it is therefore not yet equal to zero.
A Bayesian framework, expressing the notion of 'presumed error' together with the specification of a loss function, may lead to a solution.

Key input to efficiently use the above method is the presumed value p' of p. As can be seen from the chart, it is efficient for an auditor to use the upper row of the chart. In other words: if it can be assumed that the population contains no errors, a relatively small sample can be audited to efficiently assess the hypothesis that the error rate will not exceed a critical value.

2.3. Discovery Sampling

In compliance testing, the auditor only uses statistical sampling if he presumes that he can rely on the AO/IC; just to get more evidence, and to confirm this presumption, he draws a sample. As a consequence, usually one major deviation of prescribed rules implies rigorous action on behalf of the auditor (Arkin, 1984). Compliance testing in this manner can be reformulated statistically as carrying out the test:

$$H_o: p = 0 \text{ against } H_1: p > 0,$$

with acceptance number $k_o = 0$.

One of the major advantages of this specific set of hypotheses is that the probability of a type I error is zero by definition: perfect populations cannot render errors in samples. The probability of a type II error (wrongly accepting a population) is, of course, a decreasing function of the actual value of p, decreasing more steeply for larger sample sizes.

The auditor chooses a sample size by stipulating a critical value p_o and a maximally tolerated probability of a type II error β_o:

general form	Poisson	binomial
$P[\underline{k}=0 \mid N,n \text{ and } p_o] \le \beta_o$	$e^{(-np_o)} \le \beta_o$	$(1-p_o)^n \le \beta_o$
	$n \ge (-\ln \beta_o)/p_o$	$n \ge (-\ln \beta_o)/(\ln(1-p_o))$.

The choice of the parameter p_o in Discovery Sampling is made by the auditor using the notion of <u>materiality</u>. In this context, the notion of materiality is that number of errors in a population that may not pass unnoticed through an audit sample. On the problems of choosing β_o we will come back in section 4.

With both parameters chosen, the decision can also be formulated as: 'if the population error fraction exceeds p_o, the probability that the sample is nevertheless errorless may not exceed β_o'.

<u>Remark 2.1</u>
This sampling procedure is very popular amongst auditors. However, it is not uniformly most efficient in the sense of Ghosh (1970), that is the (sequential) curtailed sampling procedure (Ghosh, p. 106-107).

Kriens and Dekkers (1979) used arguments stemming from the audit approach to rephrase the general statistical testing problem, presented in section 2.2, into the problem of confirming the assumption that **certain types of errors** do not exist in a population to be audited. Their definition of these 'certain types of errors', or, as van Batenburg and Kriens (1989) call them, 'major errors' is founded on the audit strategy referred to in section 1:

> The operation of the AO/IC is tested by checking whether potential errors in a population that has been generated by the accounting system under audit have sufficiently been covered by the accounting organization and/or by measures of internal control.

In this sense, a major error is defined as a symptom of insufficient AO/IC. Therefore, the hypotheses are set in a way that only a sample with no errors will lead to acceptance of the population. The consequence of the occurrence of already one major error in the sample is a rejection of the null-hypothesis. It is, however, not a rejection of the 'fairness' of the population: it is a rejection of the assumption that the auditor can rely on the AO/IC in his verdict on the population. The auditor, therefore, will have to perform additional audit activities to assure the fairness of the population.

2.4 Error evaluation using attribute sampling

In the third phase of the auditing process, the auditor is often confronted with the following problem: there is a population of book values (trade-debts, cost invoices, purchase invoices, wages); a small portion of these values may be in error, it is assumed that all errors are positive, i.e. the book values exceed the audit values. The auditor needs an estimate for the upper bound of the total error in the population. In such a situation it is natural to compute a confidence upper bound.

Classical methods, based on a sample of entries and using the normal approximation, mostly do not lead to satisfactory results because of the small number of errors in the sample.
The earliest more satisfying solution to this problem was given by van Heerden (1961). He suggested a systematic guilder (dollar) unit sample and assigned possible errors to specific guilders in the entries. In this way a sample-guilder was either right or wrong, and so, by classical attribute sampling techniques, he could derive an upper bound for the total error amount pT in the population.

Though a great step forward as compared to classical procedures based on an evaluation of entries, the method is unsatisfactory from a statistical point of view: much information in the sample remains unused. An important improvement of the method is credited to Stringer (1963) and has become well known through publications of a.o. Leslie, Teitlebaum and Anderson (1980). With this improvement a question many auditors had struggled on was answered:

'what happens if I find an invoice on which 100 (monetary units) have been paid, but is only worth 80 ?'

In van Heerden's approach, auditors decided by the quality of the actual selected monetary unit within the item, using the convention that a possible error is located in the top monetary units of an item. In this example, there would therefore be 80% probability of finding 0, and 20% of finding 1 error. Leslie et al suggested that each monetary unit of the invoice should be considered as being '20% in error', or '20% tainted'. Their 'tainted attribute' upper confidence limit is calculated by the aid of charts. The underlying mathematical formulation is:

$$gt(k,\alpha_o)=g(0,\alpha_o) + \sum_{i=1}^{k} t(i)*[g(i,\alpha_o)-g(i-1,\alpha_o)],$$

in which $g(i,\alpha_o)$ denotes the upper limit of the $100(1-\alpha_o)$% confidence interval for $(\hat{n}p)$ using the Poisson distribution with i^o errors, and $t(i)$ the tainting of the i-th error, in a sequence from highest to lowest tainting percentage.

The use of the tainting method - often called 'Combined Attributes and Variables (CAV)-sampling' - is so widespread that it can be regarded as the standard evaluation method for auditors. To a statistician, this might be a bit surprising, realizing that:

- 25 years since the bound $gt(k,\alpha_o)$ was proposed, there has still no theoretical justification been obtained for it;
- many publications refer to simulation studies which provide strong empirical evidence of the conservativeness of the obtained upper bounds. However, only a few authors, for example Leitch et.al. (1982), actually present their simulation results;
- also this method does not exploit all information available in the sample results.

<u>Remark 2.2</u>
Cox and Snell (1979) state: 'Arguments based on treating the constituent "pounds" as independent elements seem to us (..) suspect, although the conclusions are broadly correct', and treat the sampling procedure as equivalent to random item sampling with probability proportional to size.
In our opinion their objections against monetary unit sampling are incorrect, as are the arguments by Smith (1979).

Fienberg, Neter and Leitch (1977), referring to this lack of proof of the tainting method, use a multinomial distribution approach to calculate upper bounds for the amount pT. There is a point of arbitrariness in their approach, and the calculations are tedious. Leitch et. al. (1981, 1982) cluster observations to simplify the computations; in spite of the loss of efficiency due to clustering, their method is reported to compare favorably with the bounds $gt(k,\alpha_o)$.

Quite a different approach uses Bayesian inference to incorporate prior information that is often available in audit environments (a.o. Felix and Grimlund, 1977, Cox and Snell, 1979). Just as an illustration, Cox and Snell consider a population with a small fraction phi of the book-values having a positive error. This error, expressed as a fraction of the booked value, is assumed to have an exponential distribution with parameter $1/\mu$. Next, they conceive phi and $1/\mu$ as random variables with independent Γ-prior distributions and derive the posterior distribution of psi = phi * μ. Using this distribution, an F-distribution, an upper bound for the total error psi * T can be computed. Their approach was critically evaluated by a.o. Moors (1983), Neter and Godfrey (1985), Moors and Janssens (1989), and was also part of many simulation studies.

For the time being, the overall conclusions of this section are:

- in practice, almost always the bound $gt(k,\alpha_o)$ is used;
- many alternatives have been suggested;
- there is no definite answer available to the question under what conditions which method is best.

For a more comprehensive discussion we refer to Tamura (1989), pages 13-23.

2.5 Error evaluation using variables sampling

Most important difference between attribute sampling and variables sampling is that not the sample item's quality, but its value is assessed. In auditing, this value (called the 'book value') is compared with a true value (called the 'audit value'). Furthermore, variables sampling methods are only useful once errors have been found, because these methods of estimating a 'true' population value can all be interpreted as tests whether the population mean monetary value of the errors significantly differs from 0. So, if no errors have been found, this decision has become trivial, and the sample results can only be interpreted as a SUS-sample with no errors, yielding an upper limit for the number of errors in the population of items.

As a consequence, estimation methods may lead to an approval of a population consisting of entries containing very many errors, provided their distribution is reasonably symmetrical around zero. Therefore, it is not only important that the value to be approved lies within the confidence interval, but also that this interval is relatively small.

Suppose a sample $x_1, x_2, \ldots, x_n$ of item values has been randomly taken from a population of items, and the sample mean x_d and its estimated standard error $s(x_d)$ are calculated. Assuming normality of the population item values, the t-distribution (c.q. the normal distribution) is used to construct a $100(1-\alpha_o)\%$ confidence interval for the unknown population mean μ. Multiplication by the population number of items N gives a confidence interval for the audit value of the population which can be compared to the value to be audited T.

In practice, this interval using the "direct estimator" $\underline{x}_d$ is often too wide to have any use in the auditor's decision process. Taking a larger α_o is senseless for that process, and increasing the sample size will often be too expensive, for the interval width decreases with the square root of the sample size. So, the auditor has no other choice than to use auxiliary variables to decrease the standard error of the estimator for μ. Luckily, in an audit environment there often exists an auxiliary variable with a known population mean.

In inventory audits, for example, the audit values are the sample results found by the auditor $(x_1,x_2,\ldots,x_n)$. The auxiliary variable is the book value: the n inventory values according to the client's administration $(y_1, y_2,\ldots,y_n)$ of the items corresponding to the sample items $1,2,\ldots,n$. Furthermore, the client's administration yields values N and T, so the population mean of the sample $(y_1, y_2,\ldots,y_n)$ is known and equal to T/N.

The difference estimator $\underline{x}_v$ is an unbiased estimator for the true population mean μ, and after multiplication by N an estimate for the unknown inventory value results. The formula for $\underline{x}_v$ is:

$$\underline{x}_v = \underline{x}_d + (T/N - \underline{y}_d), \text{ or}$$

$$\underline{x}_v = T/N - (\underline{y-x})_d .$$

The second formulation can be interpreted as: adjust the population mean of the auxiliary variable by the sample average of errors in that value. The first formulation says: adjust the direct estimation for the sample bias found.

In this concept, it might be possible that not necessarily 100% of this sample bias is the best way to adjust the direct estimate. To find the 'optimal adjustment' we consider the variance:

$$s^2_v = s^2(\underline{x}_d) + s^2(\underline{y}_d) - 2 \operatorname{cov}(\underline{x}_d,\underline{y}_d).$$

This expression will change if we employ an adjustment factor b*:

$$s^2_v{}' = s^2(\underline{x}_d) + b^{*2} s^2(\underline{y}_d) - 2 b^* \operatorname{cov}(\underline{x}_d,\underline{y}_d),$$

so the optimal value of $s^2_v{}'$ is when its first derivative with respect to b* is zero (and the second is positive). This implies that b* has to equal the regression coefficient b of x on y:

$$b= \operatorname{cov}(\underline{x}_d,\underline{y}_d)/s^2(\underline{y}_d) = \operatorname{cov}(\underline{x},\underline{y})/s^2(\underline{y}),$$

and the 'best' difference estimator is the regression estimator:

$$\underline{x}_r = \underline{x}_d + b (T/N - \underline{y}_d).$$

Of course, this estimator is not unbiased, but its bias is very small as long as b is close to 1. Its variance is slightly smaller than s^2_v, and very small compared to that of the direct estimator:

$$s^2_r = s^2(\underline{x}_d)\ (1 - r^2(\underline{x},\underline{y})),$$

in which $r(\underline{x},\underline{y})$ is the correlation coefficient between the x- and y values in the sample. Back to auditing, it is the correlation between the book values and the audit values of the n inventory items, so both r and b better be close to 1.

2.6 Some additional remarks

2.6.1 'SUS or MUS'

The auditor will normally choose a SUS sample for audits on non-financial errors, or on financial errors that can be either positive or negative. A SUS sample will, in attribute sampling, always result in an evaluation of the number of errors in the population of items, without any linkage to its financial consequences. In variables sampling, SUS samples are preferable because a mean monetary value per item is estimated.
A MUS sample will normally yield an evaluation of the number of errors in the population of monetary units, but only as far as positive errors (overstatements by the client) are concerned. The sample is taken from the monetary units recorded, rightly or wrongly, by the client; it can not be taken from the monetary units wrongly not recorded. Generally speaking, SUS is to be recommended in attribute sampling for audits on 'incoming flows of money', and MUS on 'outgoing flows'.

2.6.2 'Cellsampling'

Nowadays, most auditors use cellsampling, which means that the population is divided into n (equal) parts, and from each cell one random item is selected. In MUS samples, items larger than twice the cellsize will always be audited. Therefore, their evaluation can be performed separately from the other items: errors found in these items need not be extrapolated by means of interval estimation methods. Further advantages of cellsampling to the auditor are:

- the subjective notion of 'representativeness' of the sample: to the statistician, a random sample may be unevenly dispersed over a population, but the auditor feels best at ease when his sample intensity is constant over the client's book year;
- the possibility of merging, or, alternatively, division of populations: once the cellsize T/n has been determined, the population size can be increased or decreased, but the size of the sample to be audited is merely a result of counting the number of cells, and the audit perspectives as defined in monetary values are still met as specified before;
- finding the smartest fraud: as Kriens (1968) proved by formulating and solving minimax problems, the smartest fraud will minimize the

probability of being detected by dispersing his faults evenly over the population. The smartest auditor will therefore maximize this minimized probability by evenly dispersing his sample.

Cellsampling has one disadvantage: the sample is in fact not a random sample of size n, but n random samples of size 1. Hoeffding (1968) showed that the probability distribution of sample errors can still be approximated by a binomial distribution. His conclusion did not as much affect auditors, because their majority used Poisson or binomial tables, but it may have been a disappointment for an eager statistician who had programmed a hypergeometric distribution on his computer.

2.6.3.'How random is my sample ?'

Since van Heerden's systematic sample approach, most auditors have learned that a random sample is to be preferred over the risk of an unnoticed systematic error pattern. In the meantime, charts of random numbers have been replaced by random number generators (RG's). But not every calculator or PC has a good RG (Park and Miller, 1988), and not many auditors have thought about the question how to define a 'good' RG. Many audit firms have even centralized the actual production of random samples out of a standpoint of quality assessment.

Speaking for ourselves, we have found out that the majority of RG's available do not pass the 'portability test'(Knuth, 1981), implying that the same RG-computer code, supplied with exactly identical inputs, may produce different samples on different computers. For supervising and quality control this is very unattractive, because documentation of merely these inputs does not bear enough information.
Furthermore, many RG's render a scaling problem, because the random numbers produced are drawn from the range between 1 and $2^{15}-1$ and then multiplied by the ratio between population size and 32767. In fact the sample therefore only consists of multiples of this ratio: in a population size of 1.000.000, only 3.3% of the population can be represented in the sample!

We have solved both problems by implementing a 'portable' RG in LISP, using a subtractive instead of a multiplicative iteration scheme (Elsas, 1989).

3. The Average Outgoing Quality Limit method

The Average Outgoing Quality Limit sampling system was developed in the thirties by H.F. Dodge and H.G. Romig (1959). AOQL can be used to ensure that after inspection the average percentage of defectives in a series of populations will not exceed an imposed proportion p_L, by taking random samples from each separate population and testing their elements. All errors found are repaired or corrected, but if the number of errors $\underline{k}$ in a specific sample exceeds a critical value k_o, the population from which this sample was taken is rejected, implying that all its elements have to be tested and all errors have to be corrected. The value of k_o can be varied (and sample size varies consequently); its optimal value minimizes total inspection costs assuming a particular quality of the population before treatment.

AOQL was one of the various methods developed for use in the manufacture of communication apparatus and equipment for Bell Telephone Systems. The method was translated for application in an auditing environment by a.o. Cyert and Davidson (1962), Kriens and Dekkers (1979) and Arkin (1984), and for application in the control of accounting processes by Kriens and Veenstra (1986). In this setting, AOQL is successfully applied by the auditing firm Touche Ross Nederland and by a number of its clients for many years.

Unfortunately, the statistical derivation presented by Dodge and Romig is not completely correct, resulting in sample sizes that are sometimes incorrect, and sometimes in suboptimal values for k_o. In van Batenburg and Kriens (1988) some of these errors were shown, and an improved model was described.

The original version of the model by Dodge and Romig is as follows. In a population of N elements there are M errors. The number of errors in a sample of size n, denoted by $\underline{k}$, is assumed to follow a Poisson distribution with parameter np (p=M/N). If the number of errors in the sample does not exceed k_o, only these errors are corrected, otherwise all elements of the population will be tested and, if necessary, corrected. The expected value of the number of elements to be inspected is therefore:

$$I = n\ P[\underline{k} \le k_o] + N\ P[\underline{k} > k_o].$$

Dodge and Romig (mistakenly, as will be shown further on) state that the average number of errors removed equals pI. This leads them to the average fraction of errors after AOQL- treatment (p_A):

$$p_A = (M - pI)/N = p\ (N-n)/N\ P[\underline{k} \le k_o].$$

The relation between p_A and p, for given values of N, n and k_o, could be drawn to show that for small values of p, the curve is somewhat below the bisector; there is a maximum for $p = p_o$ and the curve approximates 0 for p tending to 1. If we conceive p_A as a differentiable function of p, the value of p_o can be found by equating the derivative of p_A to p equal to zero and solving for p:

$$P[\underline{k} \le k_o | np_o] - np_o\ P[\underline{k} = k_o | np_o] = 0.$$

Dodge and Romig present charts of the values of $x = np_o$ and $y = xP[\underline{k} \le k_o]$ for $k_o = 0$ to 40. The maximum value of p_A, in which p has been replaced by p_o, is p_L. Using x and y, p_L can be shown to equal:

$$p_L = (1/n - 1/N)\ y.$$

For an imposed p_L, n can now be solved as a function of N and p_L (and implicitly, a function of k_o):

$$n = Ny / (Np_L + y).$$

Because n is still a function of k_o, there are an infinite number of combinations (n,k_o) which satisfy the condition $p_A \le p_L$. Dodge and Romig tried to find the combination that minimizes the expected number I of elements to be inspected, using an estimate of the fraction of errors before treatment, p*. Even within their own model, however, their tables do not always present the optimal values, cf. Hald (1981) and Veenstra and Buysse (1985).

Taking $k_o=0$, the average number of items corrected is 0 when the population is accepted and M when it is rejected. The expected number of corrections is therefore $M\ P[\underline{k}>0]$, and not:

$$pI = np\ P[\underline{k}=0] + M\ P[\underline{k}>0].$$

Van Batenburg et al (1987) show that in the general model, the difference between the average number of corrections and pI is that between $E\{\underline{k}|\underline{k}\le k_o\}$ and $E\{\underline{k}\}$, and that ignoring this difference implies, as an ultimate consequence, that $dp_A/dp > 0$ for every $0\le p\le 1$.

So, we can conclude that Dodge and Romig's concept still holds, but their mathematical model has to be rewritten. Therefore, consider the expected outgoing quality $E[\underline{p}_A]$:

$$E[\underline{p}_A] = \sum_{k=0}^{k_o} (p-k/N)\ P[\underline{k}=k].$$

When N is 'large', k/N and n/N will be small enough to be ignored, so $\underline{k}$ can be seen as Poisson distributed and it is easy to show that the minimum value n* of the sample size that fulfils $E[\underline{p}_A] \le p_L$ for every p is:

$$n^* = y/p_L.$$

When N is 'small', $\underline{k}$ has to be interpreted as hypergeometrically distributed, and complicated computer search is necessary to find n*. However, this outcome will never exceed y/p_L. So, for small values of N taking $n^*=y/p_L$ will not be wrong, but only inefficient. A computer program, made by Philip Elsas of the Touche Ross Nederland Center for Quantitative methods and Statistics in Common Lisp, has determined the lowest value of N for which the simple formula $n^*=y/p_L$ yields the necessary sample size. If population size is below these values N*, hypergeometric calculations will yield a smaller sample size.

Chart 3.1: sample sizes from y/p_L and critical population sizes N* for different k_o and p_L.

imposed p_L	sample size from y/p_L				critical N*			
	k_o=0	k_o=1	k_o=2	k_o=3	k_o=0	k_o=1	k_o=2	k_o=3
0.5%	74	168	275	389	34657	23447	$>10^4$	$>10^4$
1%	37	84	138	195	2233	5770	$>10^5$	90984
1.5%	25	56	92	130	10453	2535	18527	20699
2%	19	42	69	98	$>10^4$	1411	6736	39602
5%	8	17	28	39	$>10^4$	300	1492	1077

Some of the results in this chart are quite surprising to us: we would expect N* to increase from left to right and to decrease from top to bottom of the table, but apparently, our subjective reasoning (or our computer program) has not yet been adequate. Furthermore, the entries '10^4', meaning that N* has not yet been found, give us some problems to think of. A possible explanation is that we have compared the hypergeometrically found sample size to the ceiling of y/p_L.

4. How to exploit available information in determining sample sizes ?

4.1 Introduction. The Audit Assurance Model

In the last few decades, the sizes of the populations to be audited have grown, resulting in a necessity to reduce the sample inaccuracy of the error fraction (to keep the inaccuracy in monetary units small enough). On the other hand, the pressure on audit costs has made a reduction of sample sizes unavoidable. Therefore, auditors and statisticians have been (re-)searching for methods that combine confidence levels the statistician can agree with, inaccuracy levels the auditor can depend on, and sample sizes the client can afford.

At the same time there was a growing feeling that determining sample sizes in the classical way - as done in most of the methods described in section 2 - is often unsatisfactory because it ignores available information about the population to be audited.

To combine these feelings with the problem mentioned in the previous paragraph, the research has been done along two completely different - and highly controversial - lines of attack: The Audit Assurance Model and Bayesian methods. We briefly review the first line and will discuss the second line in section 4.2.

The Audit Assurance Model (AAM) has appeared in many different forms. Bailey (1981) presents four, slightly different, models, with the same objective (quoted from Bailey, page 231): 'the linkage between various compliance and substantive tests of details together to render a combined reliability measure'. Each of them can be reformulated into:

$$OA = 1 - \beta_o(1-A),$$

in which:

OA= the level of <u>overall assurance</u> to be attained, the certainty that the auditor will not miss a material error (an error larger than the critical error rate p_o) in his audit;

A= the level of <u>assurance</u>, the certainty the auditor possesses that material errors will either not be present or will have been detected before the population is subjected to sampling;

β_o= the sampling risk: the probability that a population with a material error will render a sample without an error. (Because the AAM is applied within the framework of Discovery Sampling, the sampling risk is formulated as β_o instead of α_o.)

In many different versions of the AAM, the assurance (A) is divided into a number of different components, such as inherent assurance, assurance from analytical review, and assurance from compliance tests. When the American Institute was still in the first stages of discussing the Audit Assurance Model, K.A. Smith (1972) already warned:

> 'No logical basis has been determined for setting the confidence level correlated with different states of internal control. The selection of levels to be utilized is completely arbitrary, without any theoretical basis'.

We too do not believe that the AAM gives an adequate answer to the questions raised. Elsewhere we formulated our objections, based on arguments from audit theory and statistical and logical arguments (Veenstra and van Batenburg, 1989, 1990, van Batenburg, Kriens, Lammerts van Bueren and Veenstra, 1991).
Conclusion from these papers is that the AAM has been shown to be a statistically doubtful model, containing variables that should not be in it, with numerical values that can not be validated, and giving results that are methodologically not valid.

Of course, auditor's knowledge and experience, and the results of previous audit activities, may not get wasted when the auditor comes to his audit sample. Some variables in the AAM are good ways to quantify 'professional judgement'. As we have shown, the only problem is that they do not affect the confidence level used to test on a specific error fraction. They are all factors that should influence the distribution of the error fraction itself.

In Bayesian models (e.g. Kriens, 1963, Cox and Snell, 1979, Moors, 1983, van Batenburg and Kriens, 1989, 1991) the same factors can be incorporated without methodological drawbacks, and in such a way that the auditor can validate his professional judgement in monetary dimensions.

4.2 Bayesian methods

4.2.1 Introduction

The application of discovery sampling can be interpreted as the calculation of the sample size n that yields a stipulated upper limit of a $100(1-\beta_o)\%$ confidence interval for the population error fraction p if no errors have been found in the sample. When formulating such an interval, the statistician realizes that theoretically all possible values of p are 0-100%, but that an additional sample outcome will result in a somewhat smaller interval. Furthermore, a 'good' result shifts the interval towards p=0, and a 'bad' result shifts it away from p=0. Sampling can be stopped when the upper limit has descended from 100% to p_o, provided that all outcomes are 'good' ones, so the lower limit is still 0.

The number of good items it takes to bring the upper limit down to p_o does not only depend on p_o, but also on the location of this upper limit at the start of the procedure. Is it really true that without sampling the upper limit is equal to 100%?. In classical statistical theory, yes, but supported by Bayesian statistics we can start from a subjectively chosen upper limit, resulting from professional judgement and prior knowledge.

4.2.2 Bayesian Discovery Sampling

The model described in van Batenburg and Kriens (1989, 1991) starts by formulating that subjectively chosen upper limit. From that point, the sample size is calculated to derive the upper limit aimed at by the auditor. The model itself consists of:

- a Beta prior with parameters 1 (yielding a prior mode in 0, consistent with discovery sampling) and s:

  $Pr(p)= s(1-p)^{s-1}$ for $0 \le p \le 1$, $Pr(p)=0$ elsewhere;

and

- a binomial likelihood for 0 errors in a sample of size n:

  $L(k=0|p,n)= (1-p)^n$.

Together, we get a Beta posterior with parameters 1 (1 from the prior plus 0 from no errors in the sample) and s+n (s from the prior plus n from the number of errorless sample items):

$$Po(p|k=0,n)= (n+s)(1-p)^{n+s-1} \text{ for } 0 \le p \le 1, \text{ 0 elsewhere.}$$

This posterior function has to meet the auditor's requirements for discovery sampling in this year, which can now be formulated in terms of a probability of p exceeding the materiality fraction. So, the parameter (s+n) has to fulfil:

$$P(\underline{p}>p_o)= \beta_o, \text{ so } (1-p_o)^{n+s}=\beta_o, \text{ so } n+s= \log\beta_o/\log(1-p_o).$$

In the first stage, the prior parameter is derived from last year's audit results expressed as a $100(1-ß_o)\%$ upper confidence limit p* for p. In our model we replace this upper limit by the assumption:

$$P(\underline{p}>p^*)= ß^*, \text{ so } (1-p^*)^s=ß^* , \text{ so } s= \log ß^*/\log(1-p^*).$$

Some readers have suggested us to take:

$$P(\underline{p}>p^*)= ß^*(1-p^*),$$

because this probability statement can be regarded as equivalent to the formulation of the upper confidence limit on a homogeneous (Beta (1,1)) prior. We, however, do not agree on modelling 'knowing nothing about p'as a homogeneous prior, and did not follow this suggestion.

Combining the expressions for s from the prior identification and n+s, we get a sample size n_B that is sufficient for Bayesian Discovery Sampling with parameters $ß_o$ and p_o, based on the use of the prior knowledge incorporated in ß* and p*:

$$n_B = \log ß_o/\log(1-p_o) - \log ß^*/\log(1-p^*).$$

In practice, it will be rather unrealistic for the auditor to state that last year's audit sample evaluation is fully giving the right prior information for this year's prior probability function. Therefore, we incorporate a weight function f, expressing the size of the sample to be audited as a weighted average between the classically determined sample size n_C and n_B. The weight f ($0 \leq f \leq 1$) the auditor gives to his prior information is the extent to which he 'dares' to lean on his subjective prior knowledge.

Using n_C(this year)= $\log ß_o/\log(1-p_o)$ and
n_C(last year)= $\log ß^*_o/\log(1-p^*_o)$, we get:

$$n = n_C(\text{this year}) \quad - f \quad n_C(\text{last year}).$$

A numerical example, which will be referred to in section 4.2.3, shows an auditor who decides to use this year $ß_o$=0.05 and p_o=0.5% (classical binomial sample size 598), while last years' sample was 299 with ß*= 0.05 and p*=1%. (Please ignore why the auditor suddenly halves his materiality: these figures are just handy to explain the model.) Assume the auditor has taken f=40%. The BDS-sample size will be:

p_o =0.5% $ß_o$ =0.05	ß*=0.05, p* =1% s=299, f=0.40	classical sample	'Bayesian gain'	
n= 598	- 0.4 * 299	= 598	- 119	= 479

4.2.3 On the choice of the factor f

The value of f (the stability of the accounting process) is to be chosen by the auditor, quantifying the predictive power of his results from previous audit activities for this year's audit results. Therefore, we have build a three-step procedure to gather the necessary information from the audit process, and incorporated this procedure in the Touche Ross Nederland audit process UNICON and its computer assisted audit planning system COCON. This procedure is fully laid down in a manual for internal use. In short, the three steps consist of:

1. The audit expert system COCON contains a database with potential errors for each audit objective. All possible measures of AO/IC are specified, and the auditor evaluates their design, their presence and their functioning. Not every measure of internal control is necessary, but the combination of individual measures should be sufficient to give the auditor an opinion on the reliance on internal control. The database also contains audit expert's opinions on how to weigh these measures, which are called CEM (Control Evaluation Model)-scores.
   In the first step, the auditor establishes the maximum value of f based on the sum of these CEM-scores.
2. In the next step, the auditor goes through a checklist of items describing symptoms of the AO/IC that may lead to deductions of the established maximum value of f.
3. Before performing the audit sample, the auditor needs an instrument to validate the a priori added subjective information that has reduced the necessary sample size. Of course, it is impossible to validate the notion of 'weight given to prior knowledge' or 'weight given to last year's sample results'. What can be done, is validating the consequences of a specific choice of f. To show this, we use the example mentioned above. As we can see, the consequence of choosing f=0.40 is that the auditor has implicitly decided that a sample of 119 items is errorless, without having actually audited these items this year. In other words: last year's audit sample results and the evaluation of this year's AO/IC have given the auditor a 'professional judgement' that makes him (95%) sure that the error fraction in this year's population will not exceed 2.5% (the 95%-upper limit for p resulting from a sample of 119 with 0 errors). In this case the auditor's choice of f is validated by asking him:

   > 'if you want to know (with 95% certainty) whether the error fraction is below 0.5%, do you dear to lean on your advance knowledge and professional judgement that it is (with 95% certainty) below 2.5% ?'

   Of course, a specific f will not always result in the same implicitly chosen upper limit for p: this upper limit not only depends on f, but also on last year's sample size.

Apart from the individual auditor, other parties are concerned in the validation of the use of prior information. One of those is the auditor's firm, that carries out mutual quality control on the performances of individual auditors. Furthermore, the audit firm might use the collective validation of all applications to readjust the procedure leading to the 'f-values'.

In retrospect, each individual application of this method can be validated. This very argument was the most important reason for Touche Ross Nederland to replace the much-criticized Audit Assurance Model for Bayesian Discovery Sampling. In the example above, this validation can be carried out by the following reasoning. The auditor has taken a sample of 479 items, and evaluated it as if it were 598 items. If one wants to know whether this decision has been made on justified grounds, the obvious thing one can do is to audit as yet those lacking 119 items!

REFERENCES (items marked * are available on request to the authors)

Arkin, H. (1984), Handbook of Sampling for Auditing and Accounting, 3rd Ed.,New York, McGraw Hill.

Bailey, A.D. (1981), Statistical Auditing. New York, HBJ.

*van Batenburg, P.C. and J.Kriens (1988), EOQL- a revised and improved version of AOQL. Tilburg University, FEW 348.

*van Batenburg, P.C. and J.Kriens (1989), Bayesian Discovery Sampling, a simple model of Bayesian inference in auditing. The Statistician 38, 227-233

*van Batenburg, P.C. and J.Kriens (1991), Bayesian Discovery Sampling, what has happened in four years' time. Paper presented at the Fourth Valencia International Meeting on Bayesian Statistics.

*van Batenburg, P.C.,J.Kriens, W.M.Lammerts van Bueren and R.H.Veenstra (1991), Audit Assurance Model and Bayesian Discovery Sampling. Tilburg University, FEW 471.

Cox, D.R, and E.J.Snell (1979), On sampling and the estimation of rare errors. Biometrika 66, 124-132, and 69, 491 (correction).

Cyert, R.M. and H.J.Davidson (1962), Statistical Sampling for Accounting Information, Englewood Cliffs, Prentice Hall.

Dodge, H.F. and H.G.Romig (1959), Sampling Inspection Tables, 2nd Ed., New York, Wiley and Sons.

*Elsas, P.I.(1989), A Portable Random Number Generator for Application in Auditing. Touche Ross Nederland Center for Quantitative methods and Statistics; Manuscript, in Dutch.

Felix, W.L.Jr. and R.A. Grimlund (1977), Sampling Model for audit tests of composite accounts. J. Accounting Res. 15, 23-42.

Godambe, V.P. and M.E.Thompson (1988), On Single Stage Unequal Probability Sampling. In: Handbook of Statistics, 6, Sampling, Edited by P.R. Krishnaiah and C.R. Rao, Chapter 5, pp 111-123, Amsterdam, Elsevier Science Publishers.

Ghosh, B.K. (1970), Sequential Tests of Statistical Hypotheses, Reading, Addison-Wesley.

Hald, A. (1981), Statistical Theory of Sampling by Attributes, New York, Academic Press.

van Heerden, A. (1961), Statistical sampling as a means of auditing. Maandblad voor Accountancy en Bedrijfshuishoudkunde 35, 453-75 (in Dutch).

Hoeffding, W.(1956), On the distribution of the number of successes in independent trials, Annals of Mathematical Statistics, 27, 713-21.

Knuth, D.E. (1981). The art of Computer Programming, volume 2: Seminumerical Algorithms; 2nd ed., Reading, Addison-Wesley.

*Kriens, J. (1960), Random sampling in accountancy, report S 274 A/60 of the Stichting Mathematisch Centrum, Amsterdam.

*Kriens, J.(1963), De Wolff's and van Heerden's methods for random sampling in auditing, Statistica Neerlandica 17, 215-31 (in Dutch).

*Kriens, J. (1979), Statistical Sampling in Auditing, Invited paper for the 42nd Session of the International Statistical Institute. Bulletin of the I.S.I.,48, Book 3, 423-437.

*Kriens, J. (1988), Statistical Sampling in Auditing and Accounting, Tilburg University, ref.340.90.095

*Kriens, J. and A.C.Dekkers (1979), Statistical Sampling in Auditing, Leiden, Stenfert Kroese (in Dutch).
*Kriens, J. and R.H.Veenstra (1985), Statistical Sampling in Internal Control by using the AOQL-System, The Statistician 34, 383-390.
Leitch, R.A., J.Neter, R.Plante and P.Sinha (1981), Implementation of upper multinomial bounds using clustering. JASA 76, 530-533.
Leitch, R.A., J.Neter, R.Plante and P.Sinha (1982), Modified Multinomial Bounds for Larger Numbers of Errors in Audits. The Accounting Review, 57, no.2, april 1982, p 384-400.
Leslie, D.A., A.D.Teitlebaum and R.J.Anderson (1980), Dollar Unit Sampling, a practical guide for auditors, Londen, Pitman.
Moors, J.J.A.(1983), Bayes' estimation in sampling for auditing, The Statistician 32, 281-288.
Moors, J.J.A. and M.J.B.T. Janssens (1989), Exact Distributions of Bayesian Cox-Snell Bounds in Auditing. J.Acc.Res, 27, 1, Spring 1989, 135-144.
Neter, J. and J. Godfrey (1985), Robust Bayesian bounds for Monetary Unit Sampling in Auditing. Appl. Statist.34, 157-168.
Park, S.K. and K.W.Miller (1988), Random Number Generators: good ones are hard to find. Communications of the ACM, 31, 10, 1192-1201.
Roberts, D.M. (1978), Statistical Auditing, New York, American Institute of Certified Public Accountants.
Smith, K.A. (1972), the Relationship of Internal Control Evaluation and Audit Sample Size, The Accounting Review, 260-269.
Smith, T.M.F. (1979), Statistical Sampling in Auditing: a Statistician's Viewpoint. The Statistician, 28, 267-280.
Stringer, K.W. (1963), Practical aspects of statistcal sampling in auditing. Proc. Bus. Econ. Statist. Sec. 405-411. American Statistical Assocation, Washington.
Tamura, H.(1989), Statistical Models and Analysis in Auditing. Statistical Science, 4, 1, 2-33.
*Veenstra, R.H. and P.C. van Batenburg (1989, 1990), A breakthrough in statistical applications by Bayesian statistics, De Accountant 11, (July 1989), 561-564 and 1 (September 1990), 18-21 (in Dutch).
*Veenstra, R.H. and J.Buysse (1985), Optimization of the application of statistical sampling in auditing, De Accountant 91, 561-563 (in Dutch).

IN 1990 REEDS VERSCHENEN

419 Bertrand Melenberg, Rob Alessie
A method to construct moments in the multi-good life cycle consumption model

420 J. Kriens
On the differentiability of the set of efficient (μ,σ^2) combinations in the Markowitz portfolio selection method

421 Steffen Jørgensen, Peter M. Kort
Optimal dynamic investment policies under concave-convex adjustment costs

422 J.P.C. Blanc
Cyclic polling systems: limited service versus Bernoulli schedules

423 M.H.C. Paardekooper
Parallel normreducing transformations for the algebraic eigenvalue problem

424 Hans Gremmen
On the political (ir)relevance of classical customs union theory

425 Ed Nijssen
Marketingstrategie in Machtsperspectief

426 Jack P.C. Kleijnen
Regression Metamodels for Simulation with Common Random Numbers: Comparison of Techniques

427 Harry H. Tigelaar
The correlation structure of stationary bilinear processes

428 Drs. C.H. Veld en Drs. A.H.F. Verboven
De waardering van aandelenwarrants en langlopende call-opties

429 Theo van de Klundert en Anton B. van Schaik
Liquidity Constraints and the Keynesian Corridor

430 Gert Nieuwenhuis
Central limit theorems for sequences with m(n)-dependent main part

431 Hans J. Gremmen
Macro-Economic Implications of Profit Optimizing Investment Behaviour

432 J.M. Schumacher
System-Theoretic Trends in Econometrics

433 Peter M. Kort, Paul M.J.J. van Loon, Mikulás Luptacik
Optimal Dynamic Environmental Policies of a Profit Maximizing Firm

434 Raymond Gradus
Optimal Dynamic Profit Taxation: The Derivation of Feedback Stackelberg Equilibria

435 Jack P.C. Kleijnen
Statistics and Deterministic Simulation Models: Why Not?

436 M.J.G. van Eijs, R.J.M. Heuts, J.P.C. Kleijnen
Analysis and comparison of two strategies for multi-item inventory systems with joint replenishment costs

437 Jan A. Weststrate
Waiting times in a two-queue model with exhaustive and Bernoulli service

438 Alfons Daems
Typologie van non-profit organisaties

439 Drs. C.H. Veld en Drs. J. Grazell
Motieven voor de uitgifte van converteerbare obligatieleningen en warrantobligatieleningen

440 Jack P.C. Kleijnen
Sensitivity analysis of simulation experiments: regression analysis and statistical design

441 C.H. Veld en A.H.F. Verboven
De waardering van conversierechten van Nederlandse converteerbare obligaties

442 Drs. C.H. Veld en Drs. P.J.W. Duffhues
Verslaggevingsaspecten van aandelenwarrants

443 Jack P.C. Kleijnen and Ben Annink
Vector computers, Monte Carlo simulation, and regression analysis: an introduction

444 Alfons Daems
"Non-market failures": Imperfecties in de budgetsector

445 J.P.C. Blanc
The power-series algorithm applied to cyclic polling systems

446 L.W.G. Strijbosch and R.M.J. Heuts
Modelling (s,Q) inventory systems: parametric versus non-parametric approximations for the lead time demand distribution

447 Jack P.C. Kleijnen
Supercomputers for Monte Carlo simulation: cross-validation versus Rao's test in multivariate regression

448 Jack P.C. Kleijnen, Greet van Ham and Jan Rotmans
Techniques for sensitivity analysis of simulation models: a case study of the CO_2 greenhouse effect

449 Harrie A.A. Verbon and Marijn J.M. Verhoeven
Decision-making on pension schemes: expectation-formation under demographic change

450 Drs. W. Reijnders en Drs. P. Verstappen
Logistiek management marketinginstrument van de jaren negentig

451 Alfons J. Daems
Budgeting the non-profit organization
An agency theoretic approach

452 W.H. Haemers, D.G. Higman, S.A. Hobart
Strongly regular graphs induced by polarities of symmetric designs

453 M.J.G. van Eijs
Two notes on the joint replenishment problem under constant demand

454 B.B. van der Genugten
Iterated WLS using residuals for improved efficiency in the linear model with completely unknown heteroskedasticity

455 F.A. van der Duyn Schouten and S.G. Vanneste
Two Simple Control Policies for a Multicomponent Maintenance System

456 Geert J. Almekinders and Sylvester C.W. Eijffinger
Objectives and effectiveness of foreign exchange market intervention
A survey of the empirical literature

457 Saskia Oortwijn, Peter Borm, Hans Keiding and Stef Tijs
Extensions of the τ-value to NTU-games

458 Willem H. Haemers, Christopher Parker, Vera Pless and Vladimir D. Tonchev
A design and a code invariant under the simple group Co_3

459 J.P.C. Blanc
Performance evaluation of polling systems by means of the power-series algorithm

460 Leo W.G. Strijbosch, Arno G.M. van Doorne, Willem J. Selen
A simplified MOLP algorithm: The MOLP-S procedure

461 Arie Kapteyn and Aart de Zeeuw
Changing incentives for economic research in The Netherlands

462 W. Spanjers
Equilibrium with co-ordination and exchange institutions: A comment

463 Sylvester Eijffinger and Adrian van Rixtel
The Japanese financial system and monetary policy: A descriptive review

464 Hans Kremers and Dolf Talman
A new algorithm for the linear complementarity problem allowing for an arbitrary starting point

465 René van den Brink, Robert P. Gilles
A social power index for hierarchically structured populations of economic agents

IN 1991 REEDS VERSCHENEN

466 Prof.Dr. Th.C.M.J. van de Klundert - Prof.Dr. A.B.T.M. van Schaik
Economische groei in Nederland in een internationaal perspectief

467 Dr. Sylvester C.W. Eijffinger
The convergence of monetary policy - Germany and France as an example

468 E. Nijssen
Strategisch gedrag, planning en prestatie. Een inductieve studie binnen de computerbranche

469 Anne van den Nouweland, Peter Borm, Guillermo Owen and Stef Tijs
Cost allocation and communication

470 Drs. J. Grazell en Drs. C.H. Veld
Motieven voor de uitgifte van converteerbare obligatieleningen en warrant-obligatieleningen: een agency-theoretische benadering

471 P.C. van Batenburg, J. Kriens, W.M. Lammerts van Bueren and R.H. Veenstra
Audit Assurance Model and Bayesian Discovery Sampling

472 Marcel Kerkhofs
Identification and Estimation of Household Production Models

473 Robert P. Gilles, Guillermo Owen, René van den Brink
Games with Permission Structures: The Conjunctive Approach

474 Jack P.C. Kleijnen
Sensitivity Analysis of Simulation Experiments: Tutorial on Regression Analysis and Statistical Design

475 An $O(nlogn)$ algorithm for the two-machine flow shop problem with controllable machine speeds
C.P.M. van Hoesel

476 Stephan G. Vanneste
A Markov Model for Opportunity Maintenance

477 F.A. van der Duyn Schouten, M.J.G. van Eijs, R.M.J. Heuts
Coordinated replenishment systems with discount opportunities

478 A. van den Nouweland, J. Potters, S. Tijs and J. Zarzuelo
Cores and related solution concepts for multi-choice games

479 Drs. C.H. Veld
Warrant pricing: a review of theoretical and empirical research

480 E. Nijssen
De Miles and Snow-typologie: Een exploratieve studie in de meubelbranche

481 Harry G. Barkema
Are managers indeed motivated by their bonuses?

482 Jacob C. Engwerda, André C.M. Ran, Arie L. Rijkeboer
Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$

483 Peter M. Kort
A dynamic model of the firm with uncertain earnings and adjustment costs

484 Raymond H.J.M. Gradus, Peter M. Kort
Optimal taxation on profit and pollution within a macroeconomic framework

485 René van den Brink, Robert P. Gilles
Axiomatizations of the Conjunctive Permission Value for Games with Permission Structures

486 A.E. Brouwer & W.H. Haemers
The Gewirtz graph - an exercise in the theory of graph spectra

487 Pim Adang, Bertrand Melenberg
Intratemporal uncertainty in the multi-good life cycle consumption model: motivation and application

488 J.H.J. Roemen
The long term elasticity of the milk supply with respect to the milk price in the Netherlands in the period 1969-1984

489 Herbert Hamers
The Shapley-Entrance Game

490 Rezaul Kabir and Theo Vermaelen
Insider trading restrictions and the stock market

491 Piet A. Verheyen
The economic explanation of the jump of the co-state variable

492 Drs. F.L.J.W. Manders en Dr. J.A.C. de Haan
De organisatorische aspecten bij systeemontwikkeling
een beschouwing op besturing en verandering